General Relativity

Nic Ford

1   Introduction

This article is part of a series on physics for mathematicians. It is an introduction to general relativity, Einstein's famous geometric theory of gravity. One of the pleasures of learning this particular topic for mathematicians — at least the sort of mathematicians who are inclined to like geometry — is how, from the right perspective, the entire theory seems to emerge from a pretty small number of simple and geometrically plausible hypotheses. Very unlike quantum mechanics, the other big achievement of early twentieth century physics, the resulting theory is fairly easy to picture and interpret once you have the geometry under your belt.

Accordingly, the prerequisites for this piece are a bit different from many of the other articles in this series, and somewhat steeper than what's required by most textbooks on the subject. There will be a brief review in the next section, but I'm going to assume that the reader is more or less familiar with the special theory of relativity, as well as will the basics of Riemannian geometry, including the concept of a metric, the Levi-Civita connection, and the Riemann curvature tensor. (The "Connections Crash Course" article from this series might be helpful for the geometry half of this.) I'm doing this not because I expect this material to be common knowledge, but because I think this part of the story is already explained very well in a lot of other sources. I wanted to focus instead on telling the story of how the geometry turns into physics in a way that a mathematician might find pleasing, which I think is harder to get from a treatment aimed at physics students.

As usual in this series, I used a wide variety of textbooks and other resources to put this article together. Some of the ones I found most useful were:

- *General Relativity* by Robert Wald. This is the book I originally learned this topic from. It's quite comprehensive, if a bit old, and a good resource for a variety of topics and examples that are outside the scope of this article.

- *Spacetime and Geometry: An Introduction to General Relativity* by Sean Carroll. This book is a lot newer than Wald, and I think it's also easier to read. It contains a lot of nice physical motivation, and it's probably the physics book I'd recommend for a first introduction to the subject.

- *Gravitation and Cosmology: Principles and Applications of the General Theory of Relativity* by Steven Weinberg. This book is from the seventies, but I found it very clear and comprehensive; it helped me to clarify some points that I had been left wondering about from other books. Be aware that Weinberg takes a somewhat "heterodox" approach to the subject (his words!) in which he strongly downplays the geometric interpretation of the theory, more or less the opposite of what I do here.

- *General Relativity for Mathematicians* by Rainer Sachs and Hung-Hsi Wu. This book is also from the seventies. I'd also describe it as very opinionated, and I don't always agree with the authors' opinions. But, as the title suggests, it does present a lot of the material in a way that might be more palatable to mathematicians than many physics textbooks, especially when it comes to notation, and I found it to be a useful supplement to those physics textbooks even if I wouldn't recommend using it as your only source.

This article is going to be very focused on giving a clean explanation of the foundations of the theory, with only a small excursion into applications in the final section. Therefore, even more so than for the other articles in this series, I want to encourage you to check out the physics literature after you've finished with this piece; it's really in the many applications of general relativity that the beauty of the theory presents itself, so you'll be missing out on a lot of the fun if you stop before encountering more of them.

This article also has a supplement on the Lagrangian approach to general relativity. I recommend finishing this article before diving into that one.

I am grateful to Jordan Watkins and Harry Altman for helpful comments on earlier versions of this article.

2   Review and Notation

In this section, I want to briefly go over the concepts from special relativity and geometry that we're going to need, both as a refresher and to fix notation.

2.1   Special Relativity

We'll start with a lightning-fast review of special relativity. This will be in no way sufficient if you've never encountered this material before, but I hope it's a useful reminder if you have.

In special relativity, spacetime is represented by $\mathbb{R}^4$ with the **Minkowski metric**, which is an inner product of the form

$$(t, x, y, z) \cdot (t', x', y', z') = -tt' + xx' + yy' + zz'.$$

(Notice the "mostly plus" sign convention. We're using this because it is the common convention in general relativity, even though it differs from the convention used elsewhere in this series of articles! Also, we will be using units in which c, the speed of light, is equal to 1; if we weren't, there would be a c^2 multiplying the first term on the right-hand side.) As always when dealing with a metric on $\mathbb{R}^n$, we can think of the inputs to the Minkowski metric either as points in $\mathbb{R}^4$ or as tangent vectors at a single point of $\mathbb{R}^4$ depending on what's helpful in the given situation.

Given a point $p \in \mathbb{R}^4$ and a tangent vector v at p, we'll say that v is **timelike** if $v \cdot v < 0$, **spacelike** if $v \cdot v > 0$, and **lightlike** or **null** if $v \cdot v = 0$. The timelike vectors form two connected components, called the **forward-pointing** and **backward-pointing** vectors, according to whether the time component is positive or negative. The trajectory of a particle through spacetime can be represented by a smooth path $\gamma : \mathbb{R} \to \mathbb{R}^4$, sometimes called its **world-line**. If the particle is massive, the tangent vectors $\gamma'(\lambda)$ will all be timelike, which encodes the restriction that massive particles always travel slower than the speed of light. If the particle is massless (like a photon) the tangent vectors are all lightlike.

Since the time coordinate is one of the outputs of γ, nothing about the physical situation changes if we reparametrize the path on the input side. For massive particles, it's convenient to

choose a parameter τ, called **proper time**, with the property that $\gamma'(\tau) \cdot \gamma'(\tau) = -1$ everywhere. (This fixes τ up to an additive constant.) Whether we've parametrized our path with respect to proper time or not, you can compute the proper time elapsed between $\lambda = a$ and $\lambda = b$ via the arc length $\int_a^b \sqrt{-\gamma'(\lambda) \cdot \gamma'(\lambda)} d\lambda$; this quantity measures the amount of time that would be measured by a clock that's traveling alongside the particle.

If γ is parametrized with respect to proper time, the tangent vector $\gamma'(\tau)$ is called the **4-velocity** of the particle at proper time τ. The ordinary, nonrelativistic velocity of the particle then appears as the three spatial components of the 4-velocity, so the particle is at rest (according to our chosen coordinate system) if and only if its 4-velocity is $(1, 0, 0, 0)$.

If the particle has mass $m > 0$, the vector $m\gamma'(\tau)$ is called the **energy-momentum** or **4-momentum**. For massless particles, the energy-momentum is not uniquely determined by the mass and the 4-velocity but will always be a scalar multiple of the 4-velocity. As the name suggests, the energy-momentum vector carries information about what, in nonrelativistic terms, would be described as the energy and the momentum of the particle: the energy is the time component of this vector and the momentum is the three spatial components. (You might have seen some sources talk about "relativistic mass" and "rest mass," where rest mass is the m that shows up in this expression and relativistic mass is $1/c^2$ times the energy. I find this confusing, and so when I say "mass" I will always mean rest mass, which is an intrinsic property of the particle and doesn't depend on our choice of coordinates or how fast the particle is moving. Everything that could be said in terms of relativistic mass can be said using the word "energy" instead.)

The main reason to introduce the 4-velocity and energy-momentum vectors is that, unlike the concepts they're replacing (that is, velocity, energy, and momentum) they behave well under the symmetries of special relativity. Those symmetries include translations as well as the group of linear automorphisms that preserve the Minkowski metric, called $O(3, 1)$. This group has four connected components. To figure out which component an element of $O(3, 1)$ is in, you need two pieces of information: whether it preserves orientation, and whether it preserves the forward-pointing timelike vectors. The elements which preserve both (or, equivalently, which are in the connected component of the identity) are called **Lorentz transformations** and form a subgroup called $SO^+(3, 1)$. The symmetries of special relativity are the compositions of translations and Lorentz transformations.

2.2 Geometry

We now turn to a quick review of the concepts from geometry we're going to need for our later discussion. I won't assume that the reader is familiar with the notation I've chosen, especially the "abstract index notation" about to be introduced, but I will assume that the underlying concepts are ones you've seen before. To an even greater extent than in the special relativity section, this is not the place to learn these definitions for the first time!

2.2.1 Tensor Fields

Fix a smooth manifold M. Throughout the discussion we're about to have we'll need to talk a lot about sections of various vector bundles formed from tensor products of the tangent and cotangent bundles of M, so it will be useful to have some names and notation in place. Writing TM for the tangent bundle and T^*M for the cotangent bundle, we'll define the (r, s)-**tensor bundle** to be

$$T_s^r M = TM \otimes \cdots \otimes TM \otimes T^*M \otimes \cdots \otimes T^*M,$$

where there are r copies of TM and s copies of T^*M. A section of $T^r_s M$ is called an (r, s)-**tensor field** or, when it won't cause ambiguity, just an (r, s)-**tensor**. A vector field is therefore a $(1, 0)$-tensor, a 1-form is a $(0, 1)$-tensor, and a $(0, 0)$-tensor is just a real-valued function on M.

We'll have many occasions to talk about what various tensor fields look like in coordinates, so we should briefly establish some notational conventions. Given a local coordinate system $x^1, \ldots, x^n$ on some open subset $U \subseteq M$, we'll write $\partial_1, \ldots, \partial_n$ for the basis these coordinates induce on the tangent space at each point of U, and we'll write $dx^1, \ldots, dx^n$ for the basis on each of the cotangent spaces. As usual, we'll also often think of the ∂_i's and dx^i's as vector fields or 1-forms on U. Any (r, s)-tensor field on U can then be written in the form

$$\sum_{i_1=1}^{n} \cdots \sum_{i_{r+s}=1}^{n} A^{i_1 \cdots i_r}{}_{i_{r+1} \cdots i_{r+s}} (\partial_{i_1} \otimes \cdots \otimes \partial_{i_r} \otimes dx^{i_{r+1}} \otimes \cdots \otimes dx^{i_{r+s}}),$$

although thankfully we won't have much occasion to write such a hideously large expression.

2.2.2   Abstract Index Notation

Throughout this article, we'll be using a style of notation for tensor fields called **abstract index notation**. Since it's much more common among physicists than mathematicians, it's worth taking a bit of time to introduce it.

Every time we refer to an (r, s)-tensor field, its name will include a bunch of indices, r upper and s lower. For example, a vector field might be written v^a, a 1-form might be written ω_a, and a $(2, 3)$-tensor field might be written $A^{ab}{}_{cde}$. We will never write the name of a tensor without including its indices; you should think of them as part of the tensor's name.

The indices symbolize the coefficients we would have to specify in order to describe the given tensor in some local coordinate system. So, for example, the fact that we write a 1-form as ω_a is a reminder that, if we had local coordinates $x^1, \ldots, x^n$, then our 1-form could be written in the form $\omega_1 dx^1 + \cdots + \omega_n dx^n$. The reason this is called "abstract" index notation is that when we write something like ω_a we do *not* mean to imply that we have already chosen such a coordinate system, just that, if we did, those are the coefficients we'd have to specify to specify our 1-form.

We won't use the $\otimes$ symbol to denote the tensor product of two tensor fields; we'll instead just concatenate their names. For example, if v^a is a $(1, 0)$-tensor and $A^b{}_c$ is a $(1, 1)$-tensor, then $v^a A^b{}_c$ is their tensor product, which is a $(2, 1)$-tensor. This goes along very nicely with our previous interpretation of these indices: if we picked a coordinate system $x^1, \ldots, x^n$ as above and used it to write $v^a \otimes A^b{}_c$, then the coefficient of $\partial_i \otimes \partial_j \otimes dx_k$ would be $v^i A^j{}_k$.

The real utility of this notation comes from the next convention. For any point $p \in M$, there is a natural map $\kappa : T_p M \otimes T_p^* M \to \mathbb{R}$ given by $v \otimes \alpha \mapsto \alpha(v)$. Whenever the same letter appears as both an upper and lower index in an abstract index notation expression, as in $A^a{}_{ac}$, this will denote the result of applying κ to the tensor factors corresponding to that pair of indices, leaving the other tensor factors alone. This has the effect of turning an (r, s)-tensor into an $(r-1, s-1)$-tensor. In coordinates, this corresponds to summing over all possible values of the corresponding index. So if, for example,

$$A^a{}_{bc} = \sum_{i,j,k=1}^{n} A^i{}_{jk} (\partial_i \otimes dx^j \otimes dx^k),$$

then the coefficient of dx^k in $A^a{}_{ac}$ would be $\sum_{i=1}^{n} A^i{}_{ik}$.

An easy way to remember the rule is that any repeated index will always appear once up and once down (and never more than twice), and that this always means there's an implicit sum

over that index in any coordinate system. For this reason, you'll sometimes see this rule called the **Einstein summation convention**.

This process is called **contracting** the given pair of indices, and it encompasses a few different common operations on tensors. For example, if v^a is a vector field and ω_a is a 1-form, then $\omega^a v_a$ is the real-valued function you get from plugging the vector into the covector at every point. Thinking of a $(1, 1)$-tensor $A^a{}_b$ as a linear map from each tangent space to itself, the result of applying this linear map to each tangent vector in the vector field v^a is $A^a{}_b v^b$, and the real-valued function given by the trace of the linear map at each point is $A^a{}_a$.

If we happen to have chosen a local coordinate system $x^1, \ldots, x^n$, we can use it to take derivatives of tensor fields: if $A^{a_1 \cdots a_r}{}_{b_1 \cdots b_s}$ is an (r, s)-tensor, then $\partial_c A^{a_1 \cdots a_r}{}_{b_1 \cdots b_s}$ will denote the $(r, s + 1)$-tensor whose components in our chosen coordinate system are given by the corresponding partial derivatives of the components of the original tensor. This is something of an exception to a couple of our rules: unlike most uses of abstract index notation, this *does* depend on our choice of coordinate system, and it is also not meant to be interpreted as the tensor product of ∂_c with $A^{a_1 \cdots a_r}{}_{b_1 \cdots b_s}$.

2.2.3   Metrics, Connections, and Curvature

One type of tensor that will be very important to us is a **metric**, which in this language is a $(0, 2)$-tensor g_{ab} that, at every point $p \in M$, gives a nondegenerate, symmetric bilinear form on $T_p M$. In abstract index notation, the inner product of two vectors v^a and w^a is then $g_{ab} v^a w^b$. We *won't* assume the metric is positive definite, and indeed we've already seen in the Minkowski metric an example of one that is not.

You can think of a metric as a choice of isomorphism between each tangent space and its dual, i.e., as a map $T_p M \to T_p^* M$ for each $p \in M$. The inverse of this isomorphism is a $(2, 0)$-tensor which we'll call the **inverse metric** and write g^{ab}. Using these isomorphisms, it's possible to turn any (r, s)-tensor into a (p, q)-tensor as long as $r + s = p + q$. For example, we could use these isomorphisms to turn a $(2, 2)$-tensor $A^{ab}{}_{cd}$ into the $(0, 4)$-tensor $g_{ae} g_{bf} A^{ef}{}_{cd}$, or into the $(3, 1)$-tensor $g^{de} A^{ab}{}_{ce}$.

This procedure is common enough that we'll employ a notational shortcut for it, simply lower or raising an index to indicate that we're used the isomorphism arising from the metric or the inverse metric on that tensor factor. For example, we could write

$$A_{abcd} = g_{ae} g_{bf} A^{ef}{}_{cd}$$

or

$$A^{ab}{}_c{}^d = g^{de} A^{ab}{}_{ce}.$$

The ability to write all these objects succinctly is another big advantage of abstract index notation.

Given any metric g_{ab} on M, there is a unique torsion-free connection on TM which preserves g_{ab}, that is, which has the property that parallel transporting any pair of vectors preserves their inner product. This connection is called the **Levi-Civita connection**. We'll identify the connection with its **covariant derivative** operator, which we'll write ∇_a. This notation doesn't mean that ∇_a is a 1-form, rather it's an operator that takes (r, s)-tensors to $(r, s + 1)$-tensors. For example, if $A^a{}_b$ is a $(1, 1)$-tensor, then $\nabla_a A^b{}_c$ is a $(1, 2)$-tensor, and $v^a \nabla_a A^b{}_c$ is the $(1, 1)$-tensor which gives the covariant derivative of $A^b{}_c$ in the v^a direction.

Although the Levi-Civita connection is defined as a connection on the tangent bundle, you can extend it to $TM \otimes TM$ by setting

$$\nabla_a(v^b w^c) = (\nabla_a v^b)w^c + v^b(\nabla_a w^c)$$

on pure tensors, and similarly for higher tensor powers. We also get a connection on T^*M, and therefore on all the (r, s)-tensor bundles, via the requirement that

$$\omega_b(\nabla_a v^b) + (\nabla_a \omega_b)v^b = \nabla_a(\omega_b v^b);$$

because $v^b \omega_b$ is just a real-valued function, its covariant derivative has to agree with its ordinary derivative, which fixes the right-hand side of the above equation.

It will be useful for us to have a way to describe connections in coordinates. Given a coordinate system $x^1, \ldots, x^n$, the connection is completely determined by the list of numbers Γ^i_{jk} for which

$$\nabla_{\partial_j}\partial_i = \sum_{k=1}^{n} \Gamma^i_{jk}\partial_k,$$

where ∂_i denotes the i'th coordinate vector field in our chosen coordinate system. Physicists call Γ^a_{bc} the **Christoffel symbol** of the connection. (In "Connections Crash Course", this is the object we called A.)

As practice working with abstract index notation, I encourage you to verify that for an arbitrary vector field v^a we then have

$$\nabla_a v^b = \partial_a v^b + \Gamma^b_{ac}v^c.$$

Note that the left-hand side of this equation is a bona fide $(1, 1)$-tensor, but neither term on the right-hand size is a tensor. That is, for example, there is no $(1, 2)$-tensor field that, in any coordinate system $x^1, \ldots, x^n$, has Γ^i_{jk} as the coefficient on $\partial_i \otimes dx^j \otimes dx^k$. Rather, the right-hand side of this equation should just be interpreted as a recipe for how to write the coefficients $\nabla_a v^b$ in an arbitrary coordinate system. That is, it means that

$$\nabla_{\partial_j}\left(\sum_{i=1}^{n} v^i \partial_i\right) = \sum_{i=1}^{n}\left(\partial_j v^i + \sum_{k=1}^{n}\Gamma^i_{jk}v^k\right)\partial_i.$$

It's a useful exercise to work out what the corresponding expression for the covariant derivative of an arbitrary tensor field looks like.

Following the usual practice in physics, we'll write a path in M using notation like $\tau \mapsto x^a(\tau)$. This doesn't mean $x^a(\tau)$ is a vector; the notation is meant to suggest that, in any given coordinate system, we would have to specify the functions $x^1(\tau), \ldots, x^n(\tau)$ to specify the path. On the other hand, at each value of τ, the tangent vector to the path $dx^a/d\tau$ *is* a vector.

A vector field v^a is said to be **parallel transported** along a path $x^a(\tau)$ if the covariant derivative of v^a in the direction of the tangent vector to the path is always zero, that is, if

$$\frac{dx^a}{d\tau}\nabla_a v^b = 0$$

at every point along the path.

Probably the most important paths are the **geodesics**, which are the paths whose tangent vectors are parallel transported along the path itself. This is meant to capture the idea that the

path is "not accelerating," that is, the geodesics are generalizations of straight lines. I encourage you to check that this is equivalent to requiring

$$\frac{d^2 x^a}{d\tau^2} + \Gamma^a_{bc} \frac{dx^b}{d\tau} \frac{dx^c}{d\tau} = 0.$$

This is unsurprisingly called the **geodesic equation**.

Lastly, we recall the definition of **Riemann curvature tensor** $R^a{}_{bcd}$, which measures the failure of parallel transport to be path-independent or, equivalently, the failure of two different covariant derivatives to commute. For a torsion-free connection like ours, it's determined by the relation

$$[\nabla_c, \nabla_d] v^a = R^a{}_{bcd} v^b.$$

Also important are the **Ricci curvature tensor**

$$R_{ab} = R^c{}_{acb}$$

and its trace, the **scalar curvature**

$$R = R^a{}_a.$$

It will occasionally be useful to have formulas for the Christoffel symbols and the Riemann curvature tensor in coordinates. They are

$$\Gamma^a_{bc} = \frac{1}{2} g^{ad} (\partial_b g_{cd} + \partial_c g_{db} - \partial_d g_{bc})$$

and

$$R^a{}_{bcd} = \partial_c \Gamma^a_{db} - \partial_d \Gamma^a_{cb} + \Gamma^a_{ce} \Gamma^e_{db} - \Gamma^a_{de} \Gamma^e_{cb}.$$

3   Gravity as Curved Spacetime

General relativity is a relativistic theory of gravity. It can be thought of as a way of generalizing the Newtonian theory of gravity to incorporate the principles of special relativity, in particular the idea that the physically meaningful quantities are the ones that can be represented in a Lorentz-invariant way. But, as we will see, general relativity interprets those principles in a somewhat unexpected way.

Our starting point will be a simple observation about Newtonian gravity. Newton's law of gravitation says that the gravitational force exerted on a body of mass m by a body of mass M has magnitude

$$|\mathbf{F}| = \frac{GMm}{r^2},$$

where r is the distance between the two particles and G is **Newton's gravitational constant**, which is about $6.674 \times 10^{-11} \, \mathrm{N \cdot m^2 \cdot kg^{-2}}$. The force points along the line from the first body to the second. This, combined with the famous law $\mathbf{F} = m\mathbf{a}$ relating force to acceleration, produces the curious fact that the gravitational acceleration experienced by the first body doesn't actually depend on its mass at all, or in fact on anything about it other than its location.

This is not true of the other fundamental forces of nature. The electrostatic force between two charged particles, for example, is proportional to their *charges*, not their masses, while of course the relationship between force and acceleration remains the same. So the acceleration felt by a particle with charge q and mass m will depend not just on the particle's location but

also on the ratio q/m. The fact that gravity doesn't behave like this makes it unique among all of the physical interactions we know about.

If a particle is moving only under the influence of gravity and not any other forces, we will say it's **freely falling**. What our above discussion implies is that there is a distinguised family of paths in spacetime which are the ones that freely falling particles are allowed to travel along, and the path that any freely falling particle will follow is uniquely determined by its location and velocity at a single point in spacetime.

Of course, we already know of a family of paths on a certain class of manifolds which behaves exactly like this, namely the *geodesics* on a manifold with a metric. This sets us up very nicely to ask the question that will lead us to the structure of general relativity: what if the paths followed by freely falling particles *were* the geodesics with respect to some metric? This would mean that gravity is not exactly a *force* in the way that, say, electromagnetism is, since particles moving only under the influence of gravity would not be "accelerating" as such. Instead, gravity would somehow have to be encoded in the metric of spacetime itself, and this metric has to somehow depend on the distribution of matter, just as the gravitational field does in Newtonian gravity.

The idea, therefore, is that spacetime will be represented by a smooth 4-manifold M with a metric g_{ab}. Because we still expect the symmetries of special relativity to hold locally when gravity is weak, the signature of g_{ab} ought to be the same as the signature of the Minkowski metric (which for us is $(-, +, +, +)$). We'll call a metric with this signature a **pseudo-Riemannian metric**. The metric is a *dynamical variable*, that is, it depends on the physical situation rather than being fixed in advance, as it is in special relativity. Freely falling particles which are small enough not to have an appreciable effect on gravity themselves (physicists call these **test particles**) will move along geodesics.

In special relativity, as we mentioned earlier, the symmetries of spacetime are the compositions of Lorentz transformations and translations. In general relativity, by contrast, we will regard *any* diffeomorphism from our spacetime manifold M to itself as a symmetry, so long as we also transform the metric appropriately. (Wald's book notes that the contrast between these two notions of symmetry is, in fact, the source of the names "special relativity" and "general relativity.") In particular, in general relativity we lose the concept of a global "inertial reference frame." Just as, in special relativity, we expect all our laws of physics to be Lorentz-invariant, we will now want our laws of physics to be diffeomorphism-invariant.

Minkowski space, together with the accompanying framework of special relativity, will be taken to represent the situation where the effects of gravity are negligible. In the presence of gravity, though, there will not in general be any principled way to talk about rigid, global transformations of spacetime like translations, and therefore no way to directly compare things like 4-velocities or energy-momentum vectors and different points in spacetime.

At this stage, our plan for representing gravity in terms of a metric is just an intriguing idea rather than a theory — with our prescription that test particles should follow geodesics we have at least a template for how gravity ought to affect matter, but only a hazy idea of how we should describe the way matter affects gravity. In other words, we don't yet know if we can reproduce something like the predictions of Newtonian gravity using our scheme. The next couple sections will be dedicated to showing that we can.

4　The Energy-Momentum Tensor

In Newtonian gravity, the presence of mass causes a change to the *gravitational field*, which in turn exerts a force on other matter. The description of this interaction you might be most

familiar describes this effect in terms of point masses: if I have a point mass of size M located at the origin, the induced *gravitational potential* is given by

$$\Phi(\mathbf{x}) = -\frac{GM}{|\mathbf{x}|}.$$

The force felt by a particle of mass m is $\mathbf{F} = -m\nabla\Phi$. (This ∇ is a gradient, not a covariant derivative.) This means, as we alluded to earlier, that the acceleration felt by the particle is

$$-\nabla\Phi,$$

and in particular doesn't depend on any property of the particle other than its position.

Point masses are quite difficult to work with relativistically, and at any rate most of the matter in the actual universe is better described in terms of a continuous distribution of mass rather than a collection of point masses. In other words, we'd like to describe Newtonian gravity in terms of a **density function** $\rho(\mathbf{x})$. The equation that results from this is called **Poisson's equation**, and it has the rather simple form

$$\nabla^2\Phi = 4\pi G\rho.$$

(Here ∇^2 is the Laplacian, which is also often written Δ.) If ρ is spherically symmetric and supported on a ball of radius R, then there is indeed a solution to Poisson's equation for which $\Phi(\mathbf{x}) = -GM/|\mathbf{x}|$ when $|\mathbf{x}| > R$, where M is the total mass contained inside the ball.

Our goal will be to find a relativistic generalizations of these two equations. We already know more or less what's going to replace the equation describing the acceleration of particles: this is the role we've assigned to the geodesic equation. Finding our replacement for Poisson's equation is going to take a bit more work, and we'll take up that task in earnest in the next section.

Before we can do that, though, we're going to need to think a bit about what mathematical object will play the role of ρ. It might seem like we should represent it as a scalar — that is, a real-valued function, where the value at each point in spacetime represents the mass density at that point. But this turns out not to be what we want: the goal of this section is to convince you that the proper analogue of ρ in a relativistic theory is instead a $(2, 0)$-tensor field.

4.1   **Current Densities**

Using a scalar is inadequate for a couple different reasons. We can in fact see both of them in special relativity, so let's assume for the moment that we're working in Minkowski space.

The first reason is quite simple: we know from special relativity that the role played by mass in nonrelativistic physics is essentially subsumed by the concept of energy, and energy is not a scalar but the time component of the energy-momentum vector. This suggests that the relevant Lorentz-invariant quantity actually isn't mass density at all, but energy-momentum density. In other words, because we have learned from special relativity that the mass of a particle is only ever physically relevant through its contribution to energy-momentum, it is sensible to suppose that energy-momentum density, rather than just mass density, will be the quantity that serves as a source for gravity.

So why do we want a $(2, 0)$-tensor rather than a vector? The answer lies in the fact that we're trying to describe energy-momentum *density*, not energy-momentum itself. To make things a bit simpler, let's imagine for a moment that we were trying to describe the density of some

scalar quantity rather than a vector like energy-momentum. In nonrelativistic physics, this density could also be represented by a scalar. Mass density, for example, is of course mass per unit volume, and the relevant coordinate changes all preserve volumes; all coordinate systems will therefore agree about the mass density at a given point.

This is not true of the coordinate changes in special relativity, though: a Lorentz boost with velocity v will shrink volumes by a factor of $1/\sqrt{1 - v^2}$, so two different observers at the same point in spacetime can disagree about volumes, and therefore about the density of our scalar quantity at that point. But there's a standard trick in special relativity to deal with this problem.

Suppose, as will be the case for energy-momentum when we get to it below, that our scalar quantity comes attached to a continuous distribution of particles. (A good example to keep in mind is electric charge, which is indeed modeled in the way we're about to describe.) That is, there's a vector field w^a which gives, at each point in spacetime, the 4-velocity of the particle located at that point, along with a scalar function ρ_0 which gives the density of our scalar quantity as it would be measured in that particle's rest frame.

We then define the **current density** to be the vector field $j^a = \rho_0 w^a$, and I encourage you convince yourself that the time component of j^a in any coordinate system will be equal to the density that would be measured by the corresponding observer.

We can express all of this in a somewhat more coordinate-free way. For any forward-pointing unit timelike vector v^a, the density as measured by an observer whose 4-velocity is v^a will be $-v_a j^a$. If v^a is instead a spacelike unit vector, $v_a j^a$ represents the flow of our scalar quantity in the v^a direction, as it would be measured by an observer whose 4-velocity is any timelike vector orthogonal to v^a. It's often useful to unify these two concepts by thinking of density as "flow in the time direction."

With this understanding, for any unit vector v^a, $v_a j^a$ can equivalently be thought of as the *flux density* through a small 3-volume orthogonal to v^a, that is, the flux through that 3-volume divided by its volume. When v^a is spacelike, this 3-volume will look like the product of a small spacelike area with a small time interval, and so we can think of the flux as measuring how much of our scalar quantity is flowing through that area over the course of that amount of time. When v^a is timelike, the 3-volume will instead be completely spacelike, and so the flux through this 3-volume will simply give us the amount of our scalar quantity that lies inside it. (I encourage you to take some time to convince yourself of these two stories!)

A spacelike hypersurface Σ can be thought of as a "snapshot" of the universe at a moment in time. With this interpretation, the integral $\int_\Sigma (-n_a j^a)$, where n^a is the forward-pointing unit vector orthogonal to Σ at each point, represents the total amount of our scalar quantity at that moment. The divergence theorem implies that this quantity will be conserved — that is, that integral will be independent of the choice of Σ — if and only if our current density vector field is *divergence-free*.

In the standard coordinates on Minkowski space, the condition of being divergence-free can be written $\partial_a j^a = 0$. (Like most expressions involving partial derivatives, this one is *not* coordinate-free!) When this is true, we'll also say that j^a itself is a **conserved current**. I encourage you to convince yourself that, in coordinates, this equation formalizes the intuition that the only way the amount of our scalar quantity in a small region can change is if it flows into or out of the boundary of that region.

4.2   **The Energy-Momentum Tensor**

Our original question was about how to describe energy-momentum density, and energy-momentum is vector quantity, not a scalar quantity. But it is straightforward to generalize the above discussion to handle this: we'll represent energy-momentum density by a $(2, 0)$-tensor T^{ab}, so that, for any forward-pointing unit timelike vector v^a, the vector $-v_a T^{ab}$ represents the energy-momentum density that would be measured by an observer whose 4-velocity is v^a. In particular, then, the scalar $v_a v_b T^{ab}$ will be the *energy* density that that observer would measure, and in order for T^{ab} to have the physical meaning we're assigning to it this quantity will have to always be nonnegative.

We'll call T^{ab} the **energy-momentum tensor**. ("Energy-momentum current density tensor" might be more accurate, but that's understandably not a name that anyone uses. You'll also commonly see this object called the "stress-energy tensor" or even the "stress-energy-momentum tensor." All of these terms mean the same thing.)

The conversation about interpreting currents in terms of flux densities can also be adapted quite cleanly to the energy-momentum tensor. If v^a is an arbitrary vector and S is a small 3-volume orthogonal to v^a, then $v_a T^{ab}$ can be interpreted as the flux density of energy-momentum through S, where we again adopt the interpretation that "flux density" in a timelike direction should be interpreted simply as density. In particular, if we have local coordinates $(x^0, x^1, x^2, x^3) = (t, x, y, z)$ near some point in spacetime and write T^{ab} as a matrix using that coordinate system, then the (i, j) entry of that matrix is the flux density of the x^j component of energy-momentum through a small 3-surface orthogonal to the x^i direction.

In physics we expect the total energy-momentum to be a conserved quantity, and adapting the language of conserved currents we just discussed gives us a way to express this requirement in terms of T^{ab}. Given any *constant* vector field w^b, we can think of $T^{ab} w_b$ as representing the density of the component of energy-momentum in the w^b direction. (Recall that we are still in Minkowski space, which has "constant vector fields," unlike a general pseudo-Riemannian manifold!) So asking for energy-momentum to be conserved amounts to asking for each $T^{ab} w_b$ to be a conserved current, that is, for $\partial_a(T^{ab} w_b) = 0$ for all constant vector fields w^b. This will, of course, happen if and only if $\partial_a T^{ab} = 0$, so that will be the equation we use to express energy-momentum conservation.

Just as we did for currents, we can integrate T^{ab} along a spacelike hypersurface Σ to get a vector $\int_{\Sigma}(-n_a T^{ab})$ which represents the total energy-momentum in that snapshot of spacetime, and our conservation law is equivalent to asking for this vector to be independent of our choice of Σ.

We are ultimately interested not in Minkowski space but in an arbitrary 4-manifold with a pseudo-Riemannian metric. (Physicists will often refer to this change in perspective as the move to "curved spacetime.") In this context, we will definitely still want to represent energy-momentum density with a $(2, 0)$-tensor, and the interpretations we gave of the vector $-v_a T^{ab}$ and the scalar $v_a v_b T^{ab}$ when v^a is a forward-pointing unit timelike vector still stand.

The expression $\partial_a T^{ab}$ does not refer to a well-defined vector in general — that is, there is no vector field whose components look like that in every coordinate system — so our conservation law will have to take a different form. The obvious modification to make is to replace the partial derivative with a covariant derivative and say that an energy-momentum tensor is "conserved" if $\nabla_a T^{ab} = 0$.

This is indeed what we will do, but its interpretation is a bit less straightforward than it was in Minkowski space. If we were working with a current density vector field j^a and altered our definition of conserved currents in the obvious way — that is, $\nabla_a j^a = 0$ — we would be able

to adapt the story from the last section to curved spacetime more or less unchanged: if Σ is a spacelike hypersurface and ω is the volume form on Σ induced by the metric, then $\int_\Sigma \omega \cdot n_a j^a$ will still be a conserved quantity.

But a problem arises when we try to apply this philosophy to the energy-momentum tensor T^{ab}. There are no longer any "constant" vector fields w^b that we can use to extract a conserved current from T^{ab}. We might think to try integrating over a spacelike hypersurface Σ as before, but the expression $\int_\Sigma \omega \cdot n_a T^{ab}$ is no longer meaningful — it would require us to add tangent vectors that come from different points in spacetime.

There is, in fact, no way in general to associate any sort of global conservation law to the equation $\nabla_a T^{ab} = 0$. A useful way to interpret that equation is instead as a *local* conservation law. When looking at a tiny neighborhood around a single point in spacetime, it's perfectly sensible to hold onto the interpretation that "the only way the quantity of energy-momentum in this region can change is by flowing through the boundary," but there should no longer be any expectation of finding some quantity to represent the "total energy-momentum" on a spacelike hypersurface, at least not in general. (It is possible to assign a meaning to "total energy-momentum" in some special cases, and we'll discuss this problem a bit more in the final section.)

There is a useful physical perspective on this lack of global energy conservation. Our entire project involves encoding gravity in the metric on spacetime so that, in particular, gravity in general relativity is not really a "force," in the sense of arising from a potential energy function whose gradient gives the force exerted at each point. Accordingly, a phenomenon which might be described nonrelativistically as "gravity doing work on" some blob of matter will look in our theory like the energy of that blob of matter simply increasing. Put another way, T^{ab} only encodes what would be described nonrelativistically as *non-gravitational* energy and momentum, and so we should not expect it to be conserved in the same way. The difference between general relativity and Newtonian gravity, then, is that there is no good way in general to define a tensor field which represents the "energy-momentum density of the gravitational field" at each point in spacetime.

4.3 An Example

Everything we've said about the energy-momentum tensor so far pertains to the general properties an energy-momentum tensor ought to have to be physically realistic — energies should be nonnegative, so we should have $v_a v_b T^{ab} \geq 0$ for timelike v^a; and energy-momentum should be locally conserved, so we should have $\nabla_a T^{ab} = 0$. But none of this actually tells you, given some concrete physical situation you're trying to model, how to write down the corresponding energy-momentum tensor.

There are essentially two different methods for answering this question. One is to just think about how, physically, energy and momentum ought to be flowing through whatever system you're trying to describe (often using a concrete nonrelativistic model as inspiration) and try to write down a well-defined $(2,0)$-tensor field which reproduces it.

The other method, which I think it's fair to say is ultimately more robust, is to start with a Lagrangian description of your physical system. Given a Lagrangian, it's possible to derive both the form of the energy-momentum tensor for your physical system and the equation describing how the presence of matter affects gravity. (We'll be deriving this latter equation, called Einstein's equation, using a different method in the next section.)

I cover this Lagrangian perspective on general relativity in the supplement, but for the purposes of this introductory article we're going to be sticking to the first method. I thought it

would be good to mention the standard first couple of examples of energy-momentum tensors so that you can see what they tend to look like.

The first physical system we'll look at is called **dust**, which you should imagine as a collection of particles flowing through spacetime without interacting with each other. Specifically, we'll imagine that there's a timelike vector field u^a, satisfying $u_a u^a = -1$, which at each point of spacetime represents the 4-velocity of the particle located at that point. There will also be a scalar function ρ giving the mass density at each point as it would be measured by an observer travelling with the same 4-velocity as the particle, so that the 4-momentum of the particle is given by ρu^a.

I encourage you to convince yourself that, in this setup, the energy-momentum density that would be measured by an observer travelling with 4-velocity v^a is given by $\rho v_a u^a u^b$. (It will be helpful to remember that, if u^a and v^a are both 4-velocities, then $v_a u^a$ is the Lorentz factor that appears in the boost which takes one velocity to the other, and that this Lorentz factor is also the factor by which lengths are contracted in the direction of the boost.) The energy-momentum tensor of dust should therefore be

$$T^{ab} = \rho u^a u^b.$$

If you choose coordinates for the tangent space at a point in spacetime in which u^a is the vector pointing directly in the forward t direction and in which the metric looks like diag$(-1, 1, 1, 1)$, then as a matrix this looks like

$$T^{ab} = \begin{pmatrix} \rho & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

There is a slight generalization of the dust setup which appears very commonly in general relativity, and which we'll make some use of in the final section. A **perfect fluid** is a system whose energy-momentum tensor looks (under the same assumptions we used to write T^{ab} as a matrix for dust) like

$$T^{ab} = \begin{pmatrix} \rho & 0 & 0 & 0 \\ 0 & P & 0 & 0 \\ 0 & 0 & P & 0 \\ 0 & 0 & 0 & P \end{pmatrix}$$

for a scalar function P called the **pressure** of the fluid. Equivalently, in a more coordinate-free way, we can write

$$T^{ab} = (\rho + P)u^a u^b + P g^{ab}.$$

Notice that this reduces to the expression for dust when $P = 0$.

It's possible to justify this expression on physical grounds similar to the ones we used for dust, and it's also possible to derive it from a Lagrangian, but I'm going to refer you to the physics textbooks for that.

Instead, I want to close this section with a warning. From many physical theories, including Newtonian gravity, we're used to a situation where we first set up a mathematical description of whatever matter we'd like to describe, and then turn the crank of our theory to see how the relevant forces act on it. In general relativity, though, we need to know the metric in order to be able to interpret the physical content of a given energy-momentum tensor; you can see that the metric even appears explicitly in our energy-momentum tensor for a perfect fluid. What this means is that it's usually not possible to so cleanly separate the job of modeling a physical system into two steps. Instead, we usually have to solve for the metric and the energy-momentum tensor *simultaneously*, which complicates the process considerably.

$\boxed{5}$ Einstein's Equation

With the concept of the energy-momentum tensor in hand, our goal in this section is to produce the equation that describes how the presence of matter affects the metric on spacetime. While general relativity is a more general theory than Newtonian gravity and so doesn't directly follow from it, it's possible to give a fairly convincing heuristic argument for the form our equation should take using the prescription that test particles should follow geodesics and the requirement that we should reproduce Newtonian gravity in the domain where it's applicable. What follows is the version of this argument you'll find in many textbooks; in particular the presentation here was heavily inspired by Carroll's book.

We're looking for an equation which reduces to Poisson's equation in what we'll call the "nonrelativistic limit," by which we'll mean:

- The gravitational field is weak, i.e., we can write the metric in the form $g_{ab} = \eta_{ab} + h_{ab}$, where η_{ab} is the Minkowski metric and h_{ab} is sufficiently small that we can safely neglect anything beyond first order in its entries.

- The metric is *stationary*, i.e., the time derivatives of all of its entries are zero.

- All particle velocities are small compared to c. This in particular implies that all components of the energy-momentum tensor are small compared to T^{00}.

We'll be describing Newtonian gravity in terms of Poisson's equation

$$\nabla^2 \Phi = 4\pi G \rho,$$

along with our description of how particles move in the presence of gravity, which we can write in the form

$$\frac{d^2\mathbf{x}}{dt^2} = -\nabla\Phi.$$

$\boxed{5.1}$ The Geodesic Equation and Newtonian Gravity

Let's start by comparing this last equation to its general-relativistic counterpart, the geodesic equation:

$$\frac{d^2 x^a}{d\tau^2} + \Gamma^a_{bc} \frac{dx^b}{d\tau} \frac{dx^c}{d\tau} = 0.$$

(Recall that τ denotes proper time.) To examine this equation in the nonrelativistic limit, let's look at it in coordinates. We'll follow the common convention of writing either x^0 or t for the time coordinate and x^1, x^2, x^3 for the three spatial coordinates.

Our assumption that all velocities are small compared to c amounts to requiring $|dx^i/d\tau| \ll |dt/d\tau|$ for $i = 1, 2, 3$, which means we are free to only keep the term involving Γ^a_{00}. And the assumption that the gravitational field is stationary simplifies our expression for this Christoffel symbol by eliminating any terms containing a time derivative of the metric, so that we have

$$\Gamma^a_{00} = \frac{1}{2} g^{ab} (\partial_0 g_{b0} + \partial_0 g_{0b} - \partial_b g_{00}) = -\frac{1}{2} g^{ab} \partial_b g_{00}.$$

Using our assumption that $g_{ab} = \eta_{ab} + h_{ab}$, we arrive at the expression

$$\Gamma^a_{00} = -\frac{1}{2} \eta^{ab} \partial_b h_{00} + O(|h_{cd}|^2),$$

and so the geodesic equation looks like

$$\frac{d^2 x^a}{d\tau^2} - \frac{1}{2}\eta^{ab}\partial_b h_{00}\left(\frac{dt}{d\tau}\right)^2 = 0$$

to first order in the entries of h_{ab}. It's the spatial components of this that we'd like to compare to Newtonian gravity, and we can see at this point that it's quite straightforward to do so. Using the fact that all time derivatives of h_{ab} are zero, the time component of this equation just tells us that $d^2 t/d\tau^2 = 0$. This implies that, for the spatial components, $d^2 x^i/dt^2 = (d^2 x^i/d\tau^2)(dt/d\tau)^2$. (If we didn't know that $dt/d\tau$ is constant, we'd get another term from the chain rule.) So, dividing through by $(dt/d\tau)^2$ and restricting attention to the spatial components, we get

$$\frac{d^2 x^i}{dt^2} = \frac{1}{2}\partial_i h_{00}.$$

This is exactly what we needed in order to be able to make our comparison with Newtonian gravity! We'll recover our Newtonian equation if we set $h_{00} = -2\Phi$, which amounts to setting

$$g_{00} = -2\Phi - 1.$$

5.2 From Poisson's Equation to Einstein's Equation

Now that we know how the gravitational potential ought to relate to the metric in the nonrelativistic limit, we can turn our attention to Poisson's equation

$$\nabla^2 \Phi = 4\pi G\rho,$$

the other half of our description of Newtonian gravity. In our search for a relativistic version of this equation, we'll be guided by the principle that the laws of physics ought to be expressible in a way that respects the symmetries of the theory. Since, in our case, those symmetries include all local changes of coordinates, we're looking for an equation which takes the same form in all possible coordinate systems, which essentially amounts to requiring all the quantities appearing in it to be tensor fields.

Poisson's equation involves the mass density ρ, but we know that this is not a well-defined tensor relativistically; rather, mass density corresponds to the time-time component of the energy-momentum tensor, that is, to T^{00}. So, if we want a Lorentz-invariant equation, it would seem that we'll want some constant multiple of T^{ab} on the right-hand side.

The left side of Poisson's equation involves a linear combination of second derivatives of Φ, and since we know that g_{00} corresponds to $-2\Phi - 1$ in the nonrelativistic limit, we are led to look for something to put on the left-hand side of our new equation that is a linear combination of second derivatives of the components of the metric. Putting this all together, it seems that a good guess for a relativistic analogue of Poisson's equation would be something of the form

$$G_{ab} = 8\pi G T_{ab},$$

where G_{ab} is some tensor that is second order in derivatives of the metric. (Despite the unfortunate similarity in notation, the tensor G_{ab} is not directly related to G, which is still Newton's gravitational constant. The $8\pi G$ on the right side could of course be absorbed into the definition of G_{ab}; it's there both for later computational convenience and to make our definition of G_{ab} agree with the ones you'll find in the literature.)

This is actually a much stronger restriction than it might first appear. We're looking for an expression involving derivatives of the metric in which every term is second order in those derivatives, which takes the same form in every coordinate system, and which is a $(0, 2)$-tensor. (You can find an argument for this in Section 6.2 of Weinberg's book.) It turns out that the only such expressions are linear combinations of R_{ab} and $g_{ab}R$. We can therefore write

$$G_{ab} = AR_{ab} + Bg_{ab}R$$

for some scalars A and B.

We discussed in the previous section that we have physical grounds to insist that $\nabla_a T^{ab} = 0$, and so the same must be true of G_{ab}. From the contracted Bianchi identity

$$\nabla^a \left(R_{ab} - \frac{1}{2} g_{ab} R \right) = 0,$$

we learn that we need

$$\left(\frac{1}{2} A + B \right) (\nabla_a R) = 0,$$

so either R is constant or $\frac{1}{2}A + B = 0$. However, note that taking the trace of both sides of our equation for G_{ab} would give us $(A + 4B)R = 8\pi G T^a{}_a$, so if R is constant then so is $T^a{}_a$. Since it is definitely possible to have a distribution of matter which is not homogeneous in this way, we can discard this possibility and conclude that $B = -\frac{1}{2}A$.

The final degree of freedom will be taken care of by the comparison to Poisson's equation. I encourage you to check that, using the relationship between A and B just established, G_{00} looks in the nonrelativistic limit like $-A\nabla^2 g_{00}$ and that comparing this to Poisson's equation gives us that $A = 1$. (If you try to verify this, it will be helpful to take advantage of our assumption that $|T^{ij}| \ll |T^{00}|$ for $(i, j) \neq (0, 0)$ in the nonrelativistic limit.)

We therefore see that, at least if we take our starting assumptions about the form the theory ought to take seriously, we are led fairly directly to the equation

$$R_{ab} - \frac{1}{2} g_{ab} R = 8\pi G T_{ab}$$

as the relativistic analogue of Poisson's equation. This equation, called **Einstein's equation**, forms the basis for general relativity. Indeed, it's probably fair to say that general relativity amounts to little more than the claim that spacetime is a 4-manifold with a pseudo-Riemannian metric which obeys Einstein's equation. (Note also that Einstein's equation implies that T^{ab} must be *symmetric*. This isn't immediately obvious from our original definition, although it is possible to come up with some decent arguments for it on physical grounds.)

Writing $T = T^a{}_a$, it's not hard to show that Einstein's equation can also be written in the equivalent form

$$R_{ab} = 8\pi G \left(T_{ab} - \frac{1}{2} g_{ab} T \right).$$

In particular, in the absence of matter, Einstein's equation tells us that the metric has to be **Ricci-flat**, that is, the Ricci tensor must vanish. This is a strictly weaker condition than being flat, and indeed there are many nontrivial *vacuum solutions* to Einstein's equation; we'll even see one in the next section.

If we had asked for the left side of our equation to just be *at most* second order in derivatives of g_{ab} rather than exactly second order, it would also be possible to include a term proportional

to g_{ab} itself, resulting in the equation

$$R_{ab} - \frac{1}{2}g_{ab}R + \Lambda g_{ab} = 8\pi G T_{ab}.$$

(This is the only extra possibility — there are in fact no nonzero tensorial expressions at all which are exactly first order in derivatives of the metric.) The coefficient Λ is called the **cosomological constant**.

In order to have agreement with Newtonian gravity, Λ has to be very small, but there's no reason in principle why it should have to be exactly zero. Indeed, while for a long time physicists excluded the cosomological constant term, many models of the large-scale structure of the universe now include it. Notice that including this term is equivalent to adding $-\Lambda g_{ab}$ to the energy-momentum tensor. Many physicists like to do this and interpret the new term as "the energy-momentum of the vacuum" rather than as a part of the law describing how matter affects spacetime.

Einstein's equation actually contains a couple of the assumptions we used in our derivation of it. Specifically, if we hadn't used the fact that $\nabla_a T^{ab} = 0$ as part of our justification, it would follow from Einstein's equation, as would fact that test particles follow geodesics. But, beyond these two statements, Einstein's equation only covers the gravitational portion of the physical situation being modeled. If there are other, non-gravitational interactions at play — for example, if the matter that contributes to T^{ab} is interacting via electromagnetism — then Einstein's equation will have to be supplemented by some other differential equations describing those interactions.

There is a quick and dirty trick that usually tells you how to produce these equations: take whatever equation you'd use in the Minkowski-space description of the phenomenon in question and replace every ordinary derivative ∂_a with a covariant derivative ∇_a. For example, in electromagnetism we have the equation relating the field strength tensor F^{ab} to the current density J^a, which in Minkowski space takes the form $\partial_a F^{ab} = J^b$. Our trick would have us replace this with the equation $\nabla_a F^{ab} = J^b$ in curved spacetime, and that is in fact the right answer.

6 The Schwarzschild Solution

To close out this article, we're going to look at probably the simplest nontrivial solution to Einstein's equation. The discussion here is based on the versions of this story in Wald and Carroll.

The situation we'll be trying to model is a spherically symmetric universe consisting of a single body, which we'll call the "star" and model as a perfect fluid, confined to a compact region of space. We'll assume the cosmological constant is zero. We have two regions of spacetime to consider: the interior of the star, where the energy-momentum tensor will have the form we described earlier for a perfect fluid, and the exterior, where it will be zero.

6.1 The Exterior Metric

Let's start by looking at the exterior. Because $T^{ab} = 0$ in this region of spacetime, we're looking for vacuum solutions of Einstein's equation. There is a result called **Birkhoff's theorem** which strongly constrains the possible spherically symmetric vacuum solutions of Einstein's equation — there is in fact just a one-parameter family of possibilities. Writing (r, θ, ϕ) for the spherical

coordinates on $\mathbb{R}^3$, the theorem states that any spherically symmetric vacuum solution can be written in the form

$$ds^2 = -\left(1 - \frac{2GM}{r}\right)dt^2 + \left(1 - \frac{2GM}{r}\right)^{-1}dr^2 + r^2 d\Omega^2,$$

where

$$d\Omega^2 = d\theta^2 + \sin^2\theta d\phi^2$$

is the usual metric on the unit sphere in $\mathbb{R}^3$. This is called the **Schwarzschild metric with mass** M.

(Here we're employing the standard notation for writing a symmetric bilinear form on the tangent bundle in coordinates: dt^2, for example, refers to the bilinear form which takes (t, x, y, z) and (t', x', y', z') to tt', and ds^2 just refers to the metric as a whole.)

One interesting feature of the Schwarzschild metric is that it's **static**, that is, the entries are independent of t and there are no terms which mix time and space directions. Notice that this is a *conclusion* of Birkhoff's theorem, not a hypothesis.

As written here, this metric has two singularities: one at $r = 0$, and the other at the **Schwarzschild radius** $r = 2GM$. It turns out that, for ordinary astronomical bodies like the sun, the Schwarzschild radius will be well inside the object itself, and so, since the Schwarzschild metric is only meant to describe the situation outside the star, there's no problem. When this doesn't happen, we say that the star has experienced *gravitational collapse*, becoming a black hole. The resulting story is quite fun, but it's very well-covered in physics books so for now we'll be assuming that we're not in this situation.

By referring to the parameter M as "mass" we are, of course, strongly implying that it should have something to do with the mass of the star. We'll in fact go through three different ways to justify this identification, two that don't refer to the energy-momentum tensor of the star itself and one that does.

The first (and the simplest) is to look at the nonrelativistic limit we discussed when deriving Einstein's equation. In that section we saw that, in the limit as the gravitational field becomes weak, the time-time component of the metric ought to correspond to $-2\Phi - 1$, where Φ is the gravitational potential. Applying that to the Schwarzschild metric, we get $\Phi = -GM/r$, which is exactly the Newtonian gravitational potential outside a body of mass M. (In general we would only need this to happen to first order in $|g_{ab} - \eta_{ab}|$, since that was the approximation we used when comparing g_{00} to Φ before; it's a happy coincidence that we happened to get $-GM/r$ on the nose.)

6.2 Geodesics

We can arrive at a similar conclusion by looking at the geodesics in this metric, which is also an interesting thing to do in its own right. Since we have an explicit expression for the entries of g_{ab} in our chosen coordinate system, one way we could imagine proceeding is to compute all of the Christoffel symbols and write down the geodesic equations directly. I hope you will trust me, though, when I tell you that the resulting equations are hideous, and we wouldn't learn a lot from trying to attack them directly.

Luckily, another path is available to us: we can exploit the *symmetries* of the Schwarzschild metric. In pseudo-Riemannian geometry, just as in ordinary Riemannian geometry, continuous symmetries of spacetime can be captured by **Killing vector fields**, which are vector fields whose flows are isometries. While Killing vector fields don't have to exist at all in general, the

Schwarzschild metric has four linearly independent ones: one corresponding to time translation and three arising from differentiating the action of $SO(3)$, which acts by isometries due to the spherical symmetry.

If K^a is a Killing vector field and $x^a(\tau)$ is a geodesic, I encourage you to show that

$$\frac{dx^a}{d\tau} \nabla_a \left(K_b \frac{dx^b}{d\tau} \right) = 0.$$

(It will be helpful to use the fact that K^a is Killing if and only if $\nabla_a K_b + \nabla_b K_a = 0$.) In other words, $K_b (dx^b/d\tau)$ is *conserved* along the path of a geodesic. From our four linearly independent Killing vector fields we can extract four such conserved quantities, which will be very helpful for analyzing the form of the geodesics.

Because the Schwarzschild metric is preserved by a reflection through the equatorial plane (corresponding to the coordinate change $\theta \mapsto \pi - \theta$), any geodesic that starts in the equatorial plane will remain there. Since any geodesic will start in *some* plane through the origin, after an appropriate rotation we are free to assume that it is contained in the equatorial plane, i.e., that $\theta = \pi/2$.

This already does away with two of our four Killing vector fields; the remaining ones are $\partial/\partial t$ and $\partial/\partial\phi$, corresponding to time translation and rotation about the axis perpendicular to the equatorial plane. I encourage you to verify that, if $x^a(\tau) = (t(\tau), r(\tau), \theta(\tau), \phi(\tau))$ is a geodesic, then the conserved quantities we get are

$$E = \left(1 - \frac{2GM}{r} \right) \frac{dt}{d\tau}, \qquad L = r^2 \frac{d\phi}{d\tau},$$

which you can analogize to energy and angular momentum respectively.

If you're familiar with the analysis of orbits in Newtonian gravity, the second equation might look familiar: the conservation of L looks just like the one that expressed *Kelper's second law* about planets sweeping out equal areas in equal times. But this is somewhat deceptive — r is just one of the coordinates we've put on spacetime and (as is clear from looking at the coefficient of dr^2 in the metric) it *does not* represent distance from the origin. Still, it's interesting that the same formal relationship between r and $d\phi/d\tau$ appears in general relativity.

For simplicity, let's restrict our attention to timelike geodesics. (Applying this analysis to null geodesics is definitely possible, and will produce a description of how light rays bend in the presence of gravity.) If our path is timelike, then we can parametrize it in such a way that $g_{ab}(dx^a/d\tau)(dx^b/d\tau) = -1$ everywhere. Using this and the two conserved quantities just described, we can extract the following equation:

$$\frac{1}{2} \left(\frac{dr}{d\tau} \right) + \frac{1}{2} \left(1 - \frac{2GM}{r} \right) \left(\frac{L^2}{r^2} + 1 \right) = \frac{1}{2} E^2.$$

I've written this equation with an otherwise unnecessary factor of $1/2$ everywhere to highlight an interesting point: formally, this equation is identical to that of a nonrelativistic particle of energy $\frac{1}{2} E^2$ moving in a one-dimensional potential

$$V(r) = \frac{1}{2} - \frac{GM}{r} + \frac{L^2}{2r^2} - \frac{GML^2}{r^3}.$$

If you had performed this same analysis on a particle moving in a Newtonian gravitational potential, the result would have been identical except for the last term, which can be thought of as a general relativistic "correction" to the Newtonian orbit.

Of course, this correction will contribute very little if r is large, but it will start to matter more as we get closer to the star. It's possible to use this to derive one of the earliest historical successes of general relativity. Specifically, one can show that, unlike Newtonian gravity, this equation predicts orbits that are not quite elliptical. Rather, the orbits arising from solutions to this equation will *precess*, that is, the angle at which they reach their furthest point from the star will change slightly every time the geodesic completes one orbit. By the beginning of the twentieth century, even after accounting for the gravitational effects of the other planets, there was an discrepancy between theory and observation in the precession of the orbit of Mercury of about 43 arc-seconds per century, and general relativity was able to account for it quite precisely, providing an early signal that the theory was on the right track.

Examining the form of our effective potential gives us another way to justify the identification of the parameter M with the "total mass" of the star. Despite all of my warnings about how r is not actually the distance from the origin in this coordinate system, a look at the original form of the Schwarzschild metric will show that, for very large r, it is very close to the Minkowski metric in spherical coordinates. Therefore, we *are* justified in treating our coordinates more or less like ordinary flat coordinates so long as we are very far from the origin, and if we look at our potential in that context we see that it matches what we would expect from a star whose total mass is M.

A metric which, like the Schwarzschild metric, approaches the Minkowski metric at large distances is called **asymptotically flat**. Physically, you can think of an asymptotically flat space-time as representing an "isolated" system, where the matter is mostly concentrated in a finite region of space. Back in our discussion of the energy-momentum tensor, I mentioned how in general relativity there is no good way in general to talk about the "total energy-momentum" of a physical system except in special cases. Asymptotic flatness is one of those special cases — while I will refer you to the physics textbooks for the details, it's possible to perform an analysis more or less like this one to produce a definition of "total energy-momentum" in an arbitrary asymptotically flat spacetime by looking at how a test particle very far away from the origin is affected by the curvature of spacetime. (There's a nice discussion in Chapter 11 of Wald's book.)

6.3   The Interior Metric

Both of the justifications we've given so far for interpreting the parameter M as the mass of the star involved looking at what happens *outside* the star, either examining the exterior metric in the nonrelativistic limit or looking at how particles orbit around the star. While these are fairly convincing, neither one involves actually examining the energy-momentum tensor of the star itself, which is after all where one might expect to get information about the star's mass. So, as a final step, let's a look at the situation inside the star and see if we can find a relationship between its energy-momentum tensor and M.

In the interior of the star, which we decided at the start of the section to model as a perfect fluid, we're looking for solutions to Einstein's equations where

$$T^{ab} = (\rho + P)u^a u^b + P g^{ab}.$$

We'll assume that the star is *static*, in the same sense that we saw that the Schwarzschild solution is static, and that it's spherically symmetric, that is, that ρ and P are functions only of r and not of the other coordinates. This implies in particular that u^a has to be a unit vector pointing in the forward t direction.

Rather than go through the derivation of the solution here, I'll instead just state the result and point out a couple features of it. The metric that solve the equation takes the form

$$ds^2 = -e^{2a(r)}dt^2 + \left(1 - \frac{2Gm(r)}{r}\right)^{-1} dr^2 + r^2 d\Omega^2,$$

where

$$m(r) = 4\pi \int_0^r \rho(r')r'^2 dr'$$

and the function $a(r)$ solves the differential equation

$$\frac{da}{dr} = \frac{Gm(r) + 4\pi Gr^3 P}{r(r - 2Gm(r))}.$$

One of our initial assumptions was that the star is confined to some bounded region of space, so let's suppose that, for some R, the energy-momentum tensor vanishes for $r > R$. This means that, for $r > R$, our metric is the Schwarzschild metric with some mass M. In order for this transition to be continuous, we must therefore have that $M = m(R)$. In particular, we see that the energy-momentum tensor of the star does indeed determine M, which is at least somewhat comforting.

Looking at our expression for $m(r)$, it definitely seems like this relationship is a simple one: $m(R)$ looks like should be interpreted as the total mass contained in the ball of radius R. This is misleading, though: this would only be true if the volume element for our metric were what you would expect from treating (r, θ, ϕ) as spherical coordinates in flat space, and this is certainly not the case. With the right volume element in place, the total mass in the sense we're discussing now is

$$M_p = 4\pi \int_0^R \rho(r)r^2 \left(1 - \frac{2Gm(r)}{r}\right)^{-1/2} dr.$$

This quantity is sometimes called the **proper mass** of the star, and in general it will be *larger* than M, the mass which we argued can be measured by examining the star's gravitational effect on faraway particles. This gives a very nice demonstration of how sticky the concept of "total mass" can get in general relativity! The difference $M_p - M$ is usually said to be accounted for by "gravitational binding energy" of the star, and, as we've discussed, any energy-momentum that would nonrelativistically be attributed to the gravitational field itself does not show up in T^{ab}.

From our expression for the metric above and the fact that $\nabla_a T^{ab} = 0$, we can deduce that $(\rho + P)(da/dr) = -dP/dr$, and therefore that

$$\frac{dP}{dr} = -\frac{(\rho + P)(Gm(r) + 4\pi Gr^3 P)}{r(r - 2Gm(r))}.$$

This is called the **Tolman–Oppenheimer–Volkoff equation**, and gives us a relationship between $\rho(r)$ and $P(r)$. A result called **Buchdahl's theorem** proceeds from here to show that static, spherically symmetric stars of a fixed radius R must have $M < 4R/9G$. This maximum allowable mass has no analogue in Newtonian gravity.

These two issues are a good illustration of the warning that ended the section on the energy-momentum tensor: we aren't really in a position to assign a physical meaning to the components of the energy-momentum tensor until *after* we have the metric.

There is much more that we could go into here, including how we can use the Schwarzschild metric to model black holes, but I'm going to choose to end this discussion here. Even more than

for the other articles in this series, if you're interested in this topic I want to strongly encourage you to check out the physics literature on it — there are so many very fun stories that they tell much better than I ever could, and my hope is that this introduction can help you navigate that material more easily.